A WAVELET BASED CODING SCHEME VIA ATOMIC APPROXIMATION AND ADAPTIVE SAMPLING OF THE LOWEST FREQUENCY BAND

V. Bruni, D. Vitulano

Istituto per le Applicazioni del Calcolo "M. Picone", C. N. R.
Viale del Policlinico 137, 00161 Rome, Italy
Tel. +39–6–88470224, Fax +39–6–4404306
{bruni,vitulano}@iac.rm.cnr.it

ABSTRACT

In this paper[1] a new wavelet based coding scheme exploiting Wavelet Atoms and a Non Uniform Sampling (WANUS) of the lowest frequency band is presented. The encoder only sends the lowest frequency band (LFB) to the decoder. This latter exploits atomic approximation for predicting wavelet details band from LFB. Since this strategy only works in an undecimated wavelet domain, we prove here that an effective adaptive downsampling of the LFB can be reached using the minimax technique. The proposed scheme outperforms available coders in terms of rate-distortion results requiring a moderate computational effort.

Keywords: Image compression, Wavelet atoms, Adaptive sampling, Minimax technique.

1. INTRODUCTION

Image coding is gaining an increasing interest because of the variety of multimedia applications in which is involved. The actual standard JPEG2000 is based on the wavelet decomposition for compacting information [1, 2]. However, some alternative approaches have been recently proposed both in the wavelet [3] and time domain [4, 5].

In [6], a coding scheme based on the wavelet atomic decomposition has been presented. Atoms have two advantages: *i)* they allow to compact both intra and inter scale information, *ii)* they allow to link wavelet low frequency band and high frequency one at the same scale level. The coder can send just the overcomplete low frequency band (LFB), since wavelet details can be predicted from the LFB. In [6], the LFB is encoded by retaining just a few coefficients of the 2D DCT. The corresponding scheme outperforms JPEG2000 of about 1db.

In this paper, we present a new strategy. LFB is partitioned into square blocks that can contain (*edge blocks*) or not (*shade blocks*) atoms. *Edge blocks* are carefully downsampled so that high frequency prediction, i.e. atoms, is still possible. On the contrary, *shade blocks* are critically sampled. In this paper we show that it is possible to give an optimal sampling of *edge blocks* exploiting both atomic decomposition and minimax technique. Experimental results on some test images show that the proposed scheme achieves a gain up to 1.5 db over JPEG2000 rate distortion performance.

[1]This work has been supported by the Italian Ministry of Education as a part of the Firb project no. RBNE039LLC.

2. ATOMIC APPROXIMATION OF WAVELET DETAILS

In this section we present an image coding scheme that requires to only encode the approximation (lowest frequency) band of a wavelet decomposition, since wavelet details can be predicted from it. This can be achieved by the use of the atomic approximation — details can be found in [6, 8]. Wavelet coefficients of any signal can be decomposed in overlapping atoms, as the one depicted in Fig. 1. Each atom corresponds to the set of wavelet coefficients generated by the linear singularity depicted in Fig. 1.top. The atom amplitude (Fig. 1.middle) is tied to the slope of the ramp in Fig. 1.top while its shape is fixed.

One of the interesting aspects of the atomic approximation is the fact that the wavelet low frequency band can predict the high frequency one at the same scale. This can be achieved by simple linear filtering operations [6], as shortly shown below. In order to simplify the presentation, we can consider a signal with just one singularity and decompose it using the undecimated dyadic wavelet transform, i.e. $s = 2^j$, till the J^{th} scale level. If A_J indicates the LFB, let $\tilde{A}_{J-1} = A_J * \tilde{\phi}$ be the low pass component at the finer scale A_{J-1} without the detail component D_J, where $\tilde{\phi}$ is the low pass synthesis filter. Hence: $\tilde{A}_{J-1} * \phi = A_J * \tilde{\phi} * \phi$, where ϕ is the low pass analysis filter. The error for the omission of the high pass information is then $E_J = A_J - A_J * \tilde{\phi} * \phi$.

From the wavelet reconstruction formula $A_{J-1} = A_J * \tilde{\phi} + D_J * \tilde{\psi}$, where $\tilde{\psi}$ is the high pass synthesis filter, after a simple algebra we achieve:

$$E_J = A_J - A_J * \tilde{\phi} * \phi = D_J * \tilde{\psi} * \phi = \alpha_0 F_J * \tilde{\psi} * \phi, \qquad (1)$$

where $\alpha_0 F_J$ is the atom in Fig. 1.middle computed at $s = 2^J$ — it is worth outlining that the biorthogonality condition $\hat{\phi}^*(\omega)\hat{\tilde{\phi}}(\omega) + \hat{\psi}^*(\omega)\hat{\tilde{\psi}}(\omega) = 2$ guarantees that the quantities in eq. (1) are non null in the undecimated case. We can call E_J dual atom. It has the shape depicted in Fig. 1.bottom, that is very close to that of the atom in Fig. 1.middle. It is still characterized by a maximum (as in the original atom) whose location is preserved. Moreover, it still depends on the parameter α_0. It turns out that atoms locations and amplitudes can be estimated from E_J and used for recovering D_J — see [6] for details.

This operation is theoretically justified in [6], where it can be shown that atoms spread over the scale space domain through a precise law described by a partial differential equation. Mathematical details are omitted here while the algorithm is given since useful for the rest of the paper.

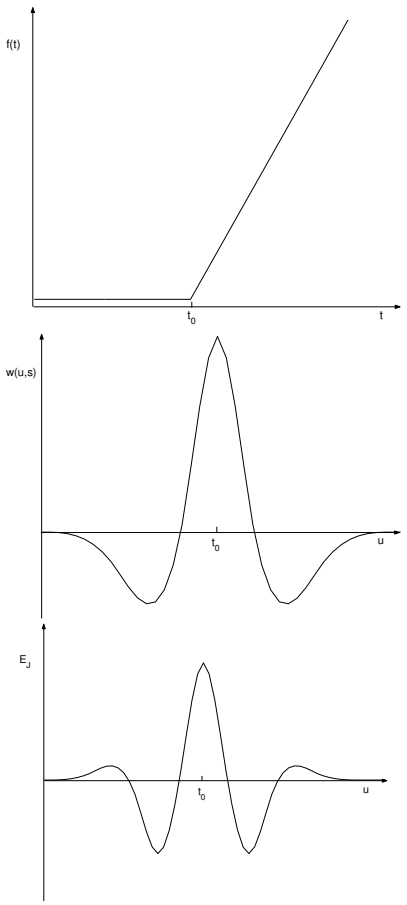

Figure 1: **(Top)** Infinite ramp signal, (**Middle**) corresponding wavelet coefficients at third scale level with the $3/9$ spline biorthogonal wavelet (basic atom), (**Bottom**) corresponding dual atom (eq. (1)).

The image is split into 1D signals. In fact, the atomic approximation works in 1D, since the tensorial product makes a distortion of the atoms shape [9].

2.1 Algorithm

Let I be an $M \times M$ image:

Coder

1. Retain and uniformly quantize the first $P \times P$ ($P \leq M$) coefficients of the 2D-DCT of the LFB A_J generated by the 1D UDWT of the rows (or columns) of I till the J^{th} scale level, using the bi-orthogonal filters $\{\Phi, \psi, \tilde{\Phi}, \tilde{\psi}\}$.

Decoder

1. Perform the inverse 2D-DCT (using zero padding) achieving an approximated LFB version $\overline{A}_J$.
2. For each row $\overline{A}_{J,r}$ of $\overline{A}_J$:
   (a) Compute $\overline{E}_J$ through the following convolution $\overline{A}_{J,r} * \psi * \tilde{\psi}$, where ψ and $\tilde{\psi}$ are respectively the high pass analysis filter and the high pass synthesis filter.
   (b) Estimate significant extrema locations $\overline{t}_k$ and their corresponding $\overline{\alpha}_k$s from the output of step 2.a — using the algorithm in [8].
   (c) Model the basic atom in Fig. 1.middle computed at $s = 2^J$ at location $\overline{t}_k$ with amplitude $\overline{\alpha}_k$ and sum each contribution to obtain the detail band $\overline{D}_J$ at J^{th} level.
   (d) Perform a step of the inverse UDWT to achieve $\overline{A}_{J-1,r}$.
3. Put $J = J - 1$ and go to step 2 till $J = 1$.

Prediction of atoms has obviously an error because of the difference between the predicted $\tilde{\alpha}_k$ and the real α_k. Therefore, the decoded image can be better recovered if the residual $\Delta\alpha_k = \alpha_k - \tilde{\alpha}_k$, is quantized and sent to the decoder as additional information.

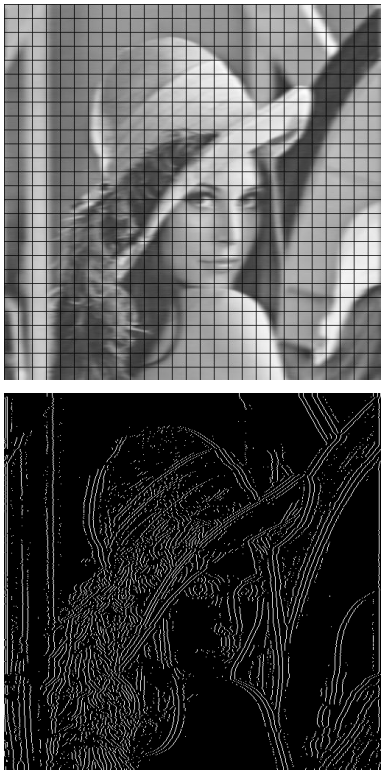

Figure 2: (**Top**) Partitioned (UDWT) LFB of Lena image at 3^{rd} scale level along rows. (**Bottom**) Corresponding significance map.

3. ADAPTIVE SAMPLING OF THE LOWEST FREQUENCY BAND

Instead of using the 2D-DCT where just a few coefficients are retained, in this section we present an alternative way for encoding the LFB. This latter can be adaptively downsampled, accounting for the atoms placement. In other words, LFB is split into square blocks, as in Fig. 2.top. They can contain at least one atom (called *edge blocks*) or not (*shade blocks*), as shown by the significance map in Fig. 2.bottom. *Shade blocks* can be critically sampled, while *edge blocks* should be carefully managed. In fact, these latter contain information that should be preserved for predicting atoms in the high frequency band. In the following we will show that it is possible to achieve an optimal sampling step using the atomic approximation and the minimax technique.

Before presenting the scheme, it is important to select the interpolant. We consider the family of wavelets generated by B-spline functions: the spline byorthogonal wavelets [7]. In the 1D case, these latter are achieved by convolving the func-

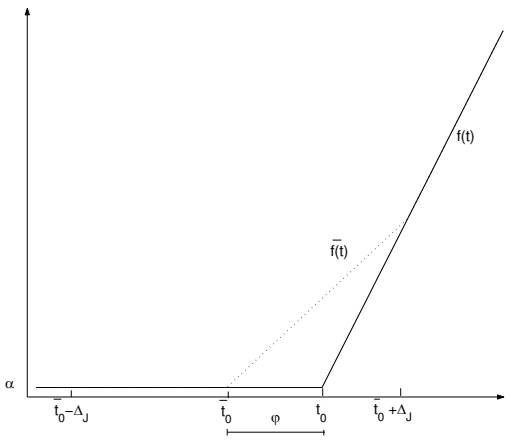

Figure 3: Original ramp function (solid line) and the recovered one (dashed) using the depicted uniform grid (step size $= \Delta_J$).

tion $\chi_{[-1,1]}(t)$ with itself n times: the order of the resulting spline is $n-1$. In particular, $n=2$ for the hat function (1^{st} order spline), while $n=4$ for the cubic B-spline. It is worth outlining that the atomic approximation is equivalent to an approximation of the original signal (in the time domain) via hat functions [10]. Hence, the LFB of a hat function using a B-spline wavelet of order m is a spline of order $m+2$. The scaling function of *rbio*2.4 wavelet (that is used in this paper) is a spline of order $m=1$; then, the LFB of a linear combination of hat functions is a linear combination of cubic splines ($m=3$). That is why we choose to interpolate the approximation band using cubic splines.

Let us consider a row of an *edge block*. Fig. 3 shows an infinite ramp f whose singularity t_0 does not coincide with any node of the grid sampling to be computed — characterized by the nodes $\bar{t}_0-\Delta_j, \bar{t}_0, \bar{t}_0+\Delta_j$. Δ_J is the step of the grid while φ measures the distance between t_0 and the previous node $\bar{t}_0$. If $Af(u,s)$ is the original LFB, $\overline{A}f(u,s)$ the LFB achieved employing the grid above and $\overline{f}$ the function (dotted line) in Fig. 3, then the distortion is: $E(u,s) = |Af(u,s) - \overline{A}f(u,s)|$. $E(u,s)$ also depends on φ and Δ_J. Obviously, the closer t_0 to the previous node, the better the recovered LFB. The minimization of $E(u,s,\varphi,\Delta_J)$ can be done via the minimax technique, i.e. $\inf_{\Delta_J}\sup_{\varphi} E(u,s,\varphi,\Delta_J)$. It can be shown (see Appendix) that the optimal sampling step $\overline{\Delta}^e_{J,r}$ at scale J along the direction r is:

$$\overline{\Delta}^e_{J,r} = \left\lfloor \frac{\overline{\Delta}_J}{2} \right\rfloor, \qquad (2)$$

where $\overline{\Delta}_J$ is the critical sampling at that scale. Along the direction c orthogonal to r, interpolation must not account for atoms prediction but just the distortion of the LFB and then

$$\overline{\Delta}^e_{J,c} = \left\lfloor \frac{\overline{\Delta}_J}{3} \right\rfloor. \qquad (3)$$

as shown in the Appendix.

As regards *shade blocks*, sampling along the c direction is equal to the critical sampling since there are no atoms.

Rows		*Columns*	
PSNR	bpp	PSNR	bpp
32.15	0.15	32.63	0.15
34.99	0.31	35.51	0.31
37.61	0.49	37.81	0.49
39.82	0.72	40.15	0.72
41.10	0.93	41.43	0.93

Table 1: PSNR and bpp on Lena image using WANUS along the horizontal and vertical direction.

Hence we have:

$$\overline{\Delta}^s_{J,r} = \overline{\Delta}_J \qquad \text{and} \qquad \overline{\Delta}^s_{J,c} = \left\lfloor \frac{\overline{\Delta}_J}{3} \right\rfloor. \qquad (4)$$

3.1 The Algorithm

We can then replace the 2D-DCT in the step 1) of both the encoder and decoder of section 2 with the adaptive sampling above.

Step 1) of **Coder**

1.a Compute significance map $T(r,c)$ — atoms locations in the detail band D_J at the coarsest scale level.

1.b Split the LFB A_J into *shade* and *edge blocks* in agreement with the criterion: at least one atom (*edge block*) or no atoms (*shade block*) — accounting for $T(r,c)$.

1.c *Edge blocks* can be sampled through $\overline{\Delta}^e_{J,r}$ and $\overline{\Delta}^e_{J,c}$ in eq. (2) and (3) respectively, while *shade blocks* through $\Delta^s_{J,r}$ and $\Delta^s_{J,c}$ in eq. (4).

1.d Samples from the previous step produce $\tilde{A}_J$, that can be DPCM along c direction and arithmetic coded.

1.e Finally, run length and entropy (arithmetic coding) code the (binary) array T_{vec} containing the blocks classification in *shade* and *edge*.

Step 1) of **Decoder**

1.a Decode T_{vec}.

1.b Decode and re-locate the samples of $\tilde{A}_J$ using T_{vec}.

1.c Interpolate $\tilde{A}_J$ along rows using the 1-D spline interpolation and then along the columns for getting the recovered LFB $\overline{A}_J$.

4. EXPERIMENTAL RESULTS

The proposed model has been tested on various images, even though we will only present here the results on $512\times512\times8$ bits Lena and Barbara images. The 1D undecimated wavelet transform (UDWT) using 3/9 rbio2.4 basis has been adopted [7]. The decomposition has been performed until the third scale level ($J=3$) both along rows and columns while the size of the square blocks (the same for both *shade* and *edge* ones) for partitioning the LFB has been selected equal to 17×17— twice the critical sampling step at the third scale.

The selected samples of the LFB have been DPCM quantized and arithmetic coded at 3.09 bpp for Lena image and 3.55 bpp for Barbara image: the quantization step has been $\Delta = \sqrt{(2)}T$, where T is the mean value of atoms amplitude at scale level J. With regard to the residual $\Delta\alpha_k$, it has been uniformly quantized and arithmetic coded at 4.5 bpp for Lena and 4.8 bpp for Barbara, with quantization step $\Delta = \frac{T}{\sqrt{2^j}}$, where j is the scale level. The vector T_{vec} has been run

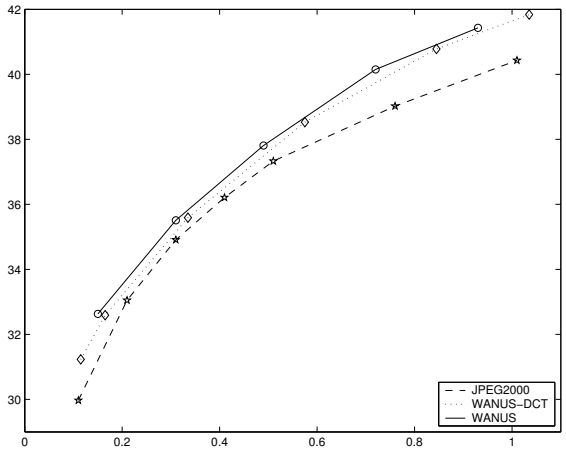

Figure 4: 512×512 bits Lena image: comparison among WANUS, JPEG2000 and the coder in [6] (WANUS-DCT).

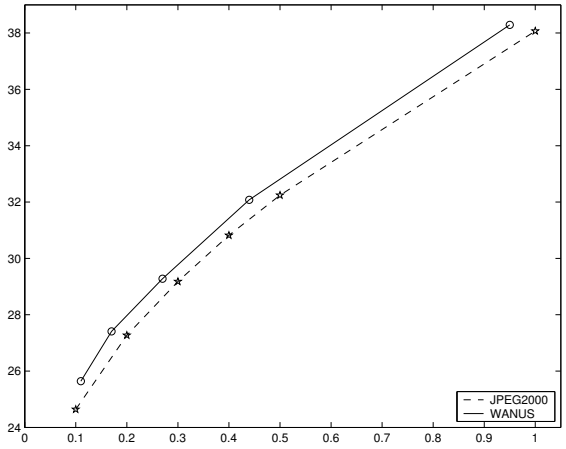

Figure 5: 512×512 bits Barbara image: comparison among WANUS and JPEG2000.

length and arithmetic coded at .62 bpp for Lena and .58 bpp for Barbara. The rate-distortion results for Lena image both using row and column direction are contained in Table 1. An interesting comparison between WANUS rate-distortion curve and WANUS-DCT in [6] (using 2D-DCT for coding the LFB) is shown in Fig. 4. Adaptive sampling gains about .5 db. WANUS has been also compared with the standard JPEG2000: the gain in terms of PSNR can reach 1.5 db at the same rate (Figs. 4 and 5).

In order to evaluate the visual quality of the recovered images, Fig. 6 shows the decoded Lena image at $bpp = 0.15$ and a zoom around its edges. As it can be observed, singularities are well reconstructed while ringing effect is reduced. A zoom of the original Barbara image and the decoded one using the proposed algorithm ($bpp = 0.17$) are also included in Fig. 5.

As regards the complexity, the coding algorithm requires: $O(NC_J)$ for the atomic decomposition, where N is the number of the atoms and C_J is the wavelet support at the scale level J; $O(\frac{M^2}{\tilde{M}^2})$ for both the classification in *shade* or *edge blocks* and for the LFB sampling, where $M \times M$ is the image size while $\tilde{M} \times \tilde{M}$ is the blocks size. The decoding algorithm requires $O(\frac{M^2}{\tilde{M}^2})$ for reorganizing the LFB data and $O(n\tilde{N})$ for their interpolation, where $n-1$ is the spline order and $\tilde{N}$ is the number of the adopted nodes.

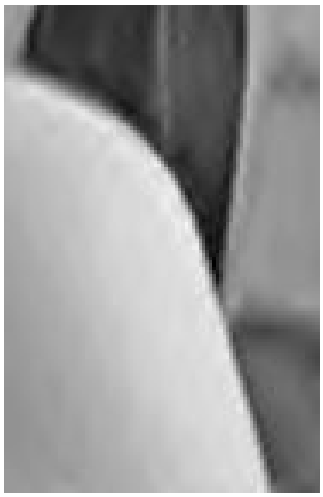

Figure 6: Decoded Lena image using WANUS at $bpp = 0.15$ and $PSNR = 32.63$ db and its detail around edges.

5. CONCLUSIONS

A novel image coding scheme called WANUS has been presented. It is based on the atomic approximation that allows to predict high frequency from the low one and an optimal non uniform sampling of the overcomplete LFB via the minimax technique. Although the intrinsic 1D nature of atomic approximation, WANUS outperforms JPEG2000 up to 1.5 db on various test images. Future developments will be oriented to try to exploit 2-D image correlation.

A. PROOF

Before applying the minimax technique for achieving the optimal sampling of both edge and shade blocks, we reorganize the distortion. In particular, exploiting linearity, shift invariance of the undecimated wavelet transform and the triangular

inequality, we have

$$E(u,s) = |Af(u,s) - \overline{A}f(u,s)| \leq \quad (5)$$
$$\leq \frac{M_\phi}{\sqrt{s}} \int_{-\infty}^{+\infty} |f(t) - \overline{f}(t)| dt,$$

where M_ϕ is the maximum of the scaling function ϕ in its support. Keeping in mind that $f(\bar{t}_0 \pm \Delta_J) = \overline{f}(\bar{t}_0 \pm \Delta_J)$ and $f(\bar{t}_0) = \overline{f}(\bar{t}_0)$, after a simple algebra we achieve:

$$|f(t) - \overline{f}(t)| = | - \alpha_0 \frac{\Delta_J - \varphi}{\Delta_J} (t - \bar{t}_0) \chi_{[\bar{t}_0, t_0]} +$$

$$+ \alpha_0 \frac{\varphi}{\Delta_J} (t - \bar{t}_0 - \Delta_J) \chi_{]t_0, \bar{t}_0 + \Delta_J]} |,$$

where $\overline{\alpha}_0 = \frac{\overline{f}(\bar{t}_0 + \Delta_J) - \overline{f}(\bar{t}_0)}{\Delta_J}$ and $\alpha_0 = \frac{f(t_0 + \Delta_J) - f(t_0)}{\Delta_J}$ and then

$$\int_{\bar{t}_0}^{\bar{t}_0 + \Delta_J} |f(t) - \overline{f}(t)| dt = |\alpha_0| \left| \frac{\Delta_J - \varphi}{2} \varphi \right|,$$

where $\overline{\alpha}_0 = \alpha_0 \frac{\Delta_j - \varphi}{\Delta_j}$. Inserting it in (5), we have:

$$E(u,s) \leq \frac{1}{\sqrt{s}} M_\phi |\alpha_0| \left| \frac{\Delta_J - \varphi}{2} \varphi \right| \equiv H(\Delta_J, \varphi). \quad (6)$$

For computing $\sup_\varphi E(\varphi, \Delta_J)$, we can now replace E with H and we use the minimax technique. We take as first estimate of Δ_J, the critical sampling $\overline{\Delta}_J$ at scale J. We then achieve $\overline{\varphi} = \frac{\overline{\Delta}_J}{2}$. Once $\overline{\varphi}$ is fixed, the new critical sampling step $\overline{\Delta}^e_{J,r}$ is the one minimizing $H(\Delta^e_J, \overline{\varphi})$ in eq. (6). It gives (2) that is the sampling step for edge blocks.

As regards shade blocks sampling, i.e. eq. (3), it comes from

$$\overline{\varphi}_c = \min \left\{ \max_\varphi \left(\int_{\bar{t}_0}^{t_0} |f(t) - \overline{f}(t)| dt \right), \right.$$
$$\left. \max_\varphi \left(\int_{t_0}^{\bar{t}_0 + \overline{\Delta}_J} |f(t) - \overline{f}(t)| dt \right) \right\} = \min \left\{ \frac{2}{3} \overline{\Delta}_J, \frac{1}{3} \overline{\Delta}_J \right\}.$$

Figure 7: *Top)* Zoom of Barbara image. *Bottom)* Zoom of the decoded Barbara image using WANUS at $bpp = 0.17$ and $PSNR = 27.41$ db.

REFERENCES

[1] A. S. Lewis, G. Knowles, *Image compression using 2-D wavelet transform*, IEEE Transactions on Image Processing, Vol. 1, pp. 244-250, February 1992.

[2] J.M. Shapiro, *Embedded image coding using zero-trees of wavelet coefficients*, IEEE Transactions on Signal Processing, Vol. 41, pp. 3445-3462, December 1993.

[3] P.L. Dragotti, M. Vetterli, *Wavelet Footprints: Theory, Algorithms and Applications*, IEEE Transactions on Signal Processing, Vol. 51, No. 5, pp. 1306-1323, May 2003.

[4] R. Shukla, P.L. Dragotti, M. Do, M. Vetterli, *Rate-Distortion Optimized Tree-Structured. Compression Algorithms for Piecewise. Polynomial Images*, IEEE Transactions on Image Processing, Vol. 14, no. 3., pp. 343-359, March 2005.

[5] R. M. Figueras i Ventura, P. Vandergheynst, P. Frossard, *Low-Rate and Flexible Image Coding with Redundand Representations*, IEEE Transactions on Image Processing, Vol. 15, No. 3, pp. 726-738, March 2006.

[6] V. Bruni, B. Piccoli, D. Vitulano, "Wavelet time-scale Dependencies for Signal and Image Compression" *Proc. of IEEE Conference ISPA 2005*, pp. 105-110, Zagreb, Croatia, September 2005.

[7] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1998.

[8] V. Bruni, D. Vitulano, Wavelet based Signal De-noising via Simple Singularities Approximation, *Signal Processing*, Vol. 86, April 2006.

[9] V. Bruni, D. Vitulano, "Combined Image Compression and Denoising using Wavelets", *Signal Processing: Image Communication*, Vol. 22 (2007), pp. 86-101.

[10] V. Bruni, B. Piccoli, D. Vitulano, "Scale Space Atoms for Signals and Image De-noising", IAC Report N. 86 (2/2006)